Translational Oncology

www.transonc.com

Volume 9 Number 4 August 2016 pp. 306–312 306

Gene Expression of the EGF System—a Prognostic Model in Non–Small Cell Lung Cancer Patients Without Activating EGFR Mutations[1]

Birgitte Sandfeld-Paulsen*, Birgitte Holst Folkersen†, Torben Riis Rasmussen†, Peter Meldgaard‡ and Boe S. Sorensen*

*Department of Clinical Biochemistry, Aarhus University Hospital, Aarhus, Denmark; †Department of Pulmonology, Aarhus University Hospital, Aarhus, Denmark; ‡Department of Oncology, Aarhus University Hospital, Aarhus, Denmark

Abstract

OBJECTIVES: Contradicting results have been demonstrated for the expression of the epidermal growth factor receptor (*EGFR*) as a prognostic marker in non–small cell lung cancer (NSCLC). The complexity of the EGF system with four interacting receptors and more than a dozen activating ligands is a likely explanation. The aim of this study is to demonstrate that the combined network of receptors and ligands from the EGF system is a prognostic marker. *MATERIAL AND METHODS:* Gene expression of the receptors *EGFR*, *HER2*, *HER3*, *HER4*, and the ligands *AREG*, *HB-EGF*, *EPI*, *TGF-α*, and *EGF* was measured by quantitative polymerase chain reaction in tumor samples from 100 NSCLC patients without EGFR activating mutations. Results were dichotomized into high or low levels of target expression. Coexpression of the ligands and receptors was observed, and a score was developed based on the summed effect of receptors and ligands. Akaike's information criteria selected the optimal score. Results were correlated with age, sex, stage, histology, performance status, and overall survival. *RESULTS:* Patients were randomly split 1:1 to create test and validation cohorts. In multivariate analyses, the only individual prognostic marker was *EPI* (hazard ratio [HR] 0.38 [0.20-0.72], $P = .003$). The optimal score in the test cohort was validated as a marker of inferior survival in the validation cohort and by bootstrapping. Multivariate analysis confirmed the combined score as a prognostic marker of inferior survival (HR 3.75 [2.17-6.47], $P < .00001$). *CONCLUSION:* Our study has developed a model that takes the complexity of the EGF system into account and shows that this model is a strong prognostic marker in NSCLC patients.

Translational Oncology (2016) 9, 306–312

Despite advances in the treatment, non–small cell lung cancer (NSCLC) remains the leading cause of cancer-related death in the world [1]. In particular, the overall prognosis is poor for the metastatic stages, with a median overall survival (OS) of only 8 to 10 months. Even in the early nonmetastatic stages, the 5-year survival rate is as low as 50% [2,3]. Prognostic markers are needed to stratify patients with different risk outcome. Several biomarkers have been evaluated in NSCLC, but only a few have proven to be clinically relevant. An activating mutation in the epidermal growth factor receptor (*EGFR*) is both a well-described predictive marker of benefit of EGFR-targeted tyrosine kinase inhibitors but also a debated prognostic marker of better OS [4–9]. As *EGFR* expression has been associated with OS in head and neck, colorectal, and esophagus cancer [10–12], attention has been directed toward the use of *EGFR* expression as a prognostic marker in NSCLC, but contradicting results have been demonstrated [13–16]. The EGF system is complex, and the effect of ligand-receptor interaction depends on a variety of different factors, which provides a plausible explanation for the divergence observed between studies that only evaluate *EGFR* expression in general. *EGFR* is one out of four related receptors from the EGF system and is

Address all correspondence to: Birgitte Sandfeld Paulsen, Department Of Clinical Biochemistry, Aarhus University Hospital, 8000 Aarhus, Denmark.
E-mail: birgne@rm.dk

[1] Conflict of interest: The authors have no conflict of interest, and this study had no specific funding.

Received 6 April 2016; Accepted 9 May 2016

1936-5233/16
http://dx.doi.org/10.1016/j.tranon.2016.05.002

capable of forming homodimers or heterodimers with one of the three other receptors when activated by a ligand. Several ligands from the EGF system such as amphiregulin (AREG), epidermal growth factor (EGF), and transforming growth factor–α (TGF-α) only activate EGFR, whereas some have the ability to activate several combinations of the four EGF receptors like heparin-binding epidermal growth factor (HB-EGF), epiregulin (EPI), and betacellulin (BCL). Most knowledge on the role of the ligands in NSCLC is from *in vitro* studies or from smaller clinical studies. *In vitro* studies have suggested that the biological effect of EGFR activation is dependent on the specific activating ligand as well as the dimerization partner [17]. Yet, no clinical studies have evaluated the effect of the network of receptors and ligands influencing EGFR in NSCLC. Furthermore, the majority of the clinical studies exploring *EGFR* expression are based on immunohistochemistry which is a semiquantitative method with a great risk of interobserver variability. Quantitative gene expression analyses provide a more accurate measure and are therefore more suitable for studies comparing expression levels. Prospectively, we have collected fresh tumor samples from patients suspected of lung cancer. Accordingly, the aim of this study is to evaluate the gene expression of the network of receptors and ligands of the EGF system affecting EGFR as a prognostic markers in NSCLC.

1. Materials and Methods

1.1. Patients and Tumor Samples

In this study, 1093 patients referred to the Department of Pulmonary Medicine, Aarhus University Hospital, Denmark, for diagnostic workup of lung cancer were included consecutively from April 2011 until January 2013. Patients with NSCLC without mutations in *EGFR* defined the NSCLC group used in this study. Patients with other types of cancer than NSCLC were excluded. To establish normal values of gene expression, a reference group was made of patients without cancer, as they had comparable clinical features (age, smoking history, clinical symptoms, and anamnestic symptoms leading to lung cancer suspicion). Patient selection is depicted in Figure 1. Clinicopathological characteristics were collected at time of inclusion. To minimize the risk of positive overestimation when establishing a score, the NSCLC group was randomly split 1:1 into a test and a validation cohort matched on histology, stage, WHO performance status (PS), and sex.

All patients gave informed written consent before inclusion, and the study was approved by the Central Denmark Region Committees on Biomedical Research Ethics (M-20100246).

Tumor samples were collected at time of diagnosis with the Ultra-micro sampling technique, allowing an aliquot of the diagnostic sample to be used for gene expression studies [18]. Samples were stored in RNAprotect at −80°C until total RNA extraction.

1.2. RNA Extraction and cDNA Preparation

Total RNA was extracted from all tissue samples using the RNeasy micro kit from Qiagen, (Qiagen, Hilden, Germany; www.qiagen.com) according to the instructions of the manufacturer. Purified RNA was diluted in 14 μl of RNase-free H_2O. cDNA was synthesized from 2 μl of total RNA in a 20-μl reaction, following manufacturer's instructions (Applied Biosystems Inc., CA).

1.3. Quantitative Polymerase Chain Reaction (PCR)

Expression of *EGFR*, *HER2*, *HER3*, *HER4*, *AREG*, *HB-EGF*, *EPI*, *TGF-α*, *EGF*, and the reference gene *β-actin* was quantified using the LightCycler 480 instrument. The PCR was performed using SYBR Green I Master Mix Buffer except HER2 that was analyzed by hydrolysis probes (SYBR Green RT-PCR Reagents Kit, Roche, Mannheim, Germany) following manufacturer's instructions. Primer sequences, concentrations, and annealing temperature are shown in Table S1. All primers were purchased from Eurofins MWG Synthesis GmbH, Ebersberg, Germany. In an earlier study, we found that, in these samples, *β-actin* was the most stable household gene [18]. Therefore, each target gene was normalized to the household gene *β-actin* and used for further analysis.

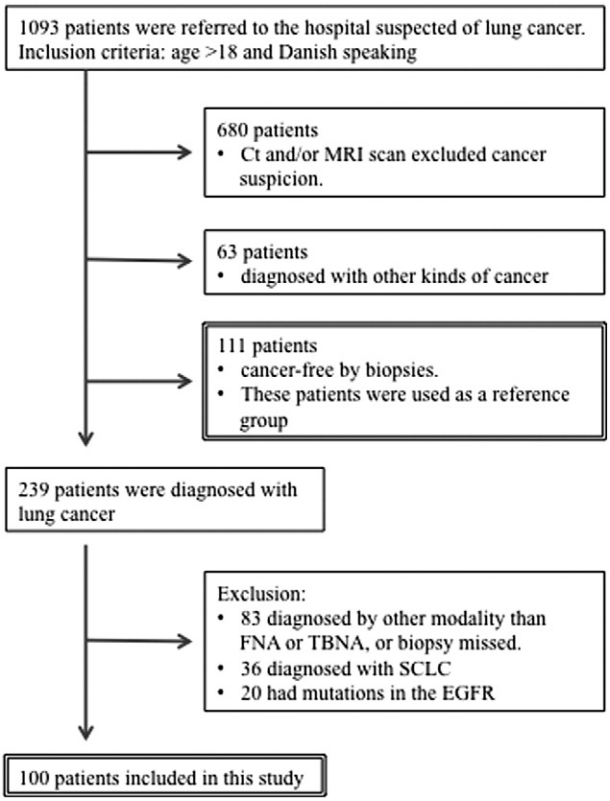

Figure 1. Patient selection.

1.4. Cut Points

A cut point was determined in the reference group for each target gene. The cut point was defined by the mean + 2 SD for the individual target gene, and expression levels above the cut points were defined as elevated.

1.5. The EGF Score Model

To evaluate the EGF system as a network, a score based on the summed effect of each receptor and ligand was established. Each ligand and receptor were given the value of +1 or −1 based on the individual markers' hazard ratio (HR) in the univariate analysis in the test cohort [19]. If the HR was greater than 1, the value given was +1, and if the HR was less than 1, the value was −1. The sum of all markers created the score. The score was then confirmed by bootstrapping (1000 times) to create bias-corrected confidence intervals (BCIs). Redundancy of one or several markers was tested to optimize the score. To select the score with the best model fit to explain the data, Akaike's information criteria (AIC) were calculated in the test cohort for every combination excluding one or several of the markers. When the difference of the AIC values was larger than 2,

Table 1. Clinical Characteristics

	Test Cohort $n = 50$	Validation Cohort $n = 50$	P Values*,†	Reference Cohort n = 111 (282 Samples)	P Values*,‡
Age	66.4	68.7	.21	60.7	<.001
mean (min-max)	(37-86)	(50-86)		(21-84)	
Sex *n* (%)					
Female	30 (60)	20 (40)	.05	51 (46)	.51
Male	20 (40)	30 (60)		60 (54)	
Smoking					
Never	2	1	.23	28	<.001
Former	26	32		59	
Current	22	17		24	
PS *n* (%)					
0	21 (42)	15 (30)	.24		
1	17 (34)	27 (54)			
2	10 (20)	7 (14)			
3	2 (4)	1 (3)			
Stage					
I & II	11 (22)	9 (18)	.44		
III & IV	39 (78)	41 (82)			
Histology					
Squamous	15 (30)	12 (24)	.50		
Nonsquamous	35 (70)	38 (76)			
Sample origin					
Tumor	18 (36)	11	.12	34	.02
Lymph node	32 (64)	39		248	

* P values calculated by Person χ^2 or t test.
† Test cohort versus validation cohort.
‡ Reference cohort versus test and validation cohorts.

the score with the lowest AIC was selected. If the difference of the AIC was less than 2, the score with the least target genes was selected [20]. Based on the calculated AIC, the score with the best model fit in the test cohort was chosen and validated in the validation cohort. By bootstrapping (1000 times), the optimal score was validated. After validation, a multivariate analysis was carried out in the entire NSCLC group taking clinicopathological characteristics into account.

1.6. Statistical Analysis

Associations between clinical characteristics were tested by Person χ^2 or t test. Survival was calculated from date of diagnosis until date of death or last day of follow-up (May 1, 2015). OS was defined as death of any cause or last date of follow-up. OS was estimated by the Kaplan-Meier method and compared by the log-rank test. Uni- and multivariate HRs were estimated by Cox proportional hazard analyses. HR is presented with a 95% CI. All P values are two-sided with a 5% significance level. The proportional hazard assumption was assessed graphically by plotting Schoenfeld residuals and log-log plots. Bootstrapping was used for validation of estimated HR. Bias in the bootstrap-derived calculations was used as an indicator of robustness of estimates. To identify effect modification, interaction terms were established and added to the regression model. Backward selection was used to identify the multivariate model with best fit. Variables were stepwise removed until only significant variables remained ($P < .10$). Model fit was evaluated by AIC. Statistical analyses were carried out in STATA version 13.

2. Results

2.1. Patient Characteristics

Clinical and pathological characteristics are detailed in Table 1. One hundred patients with NSCLC without *EGFR* mutations and 111 patients without cancer were included. From the NSCLC patients, a test cohort ($n = 50$) was randomly generated with a similar distribution of age, stage, PS, and histologic subtype as the total cohort. In the reference cohort, the mean age was lower (60.7) than in the cancer cohorts (67.5) and the distribution of never smokers versus current or former was significantly different. Furthermore, in the reference cohort, more samples originated from lymph nodes compared with the primary tumor than in the cancer cohorts.

2.2. The EGF System

Reference group: 282 samples each representing individual lesions from 111 patients were examined, and a cut point for each target gene was calculated as described above. Because this part of the study was carried out to characterize the background expression of the target genes, all available samples from each patient were included. Cut points are shown in Table 2.

NSCLC group: Level and distribution of each target gene are shown in Figures 2 and S1 and in Table 2. For the receptors *EGFR* and *HER3* and the ligands, elevated levels were observed in both nonsquamous and squamous cell type and predominantly in the advanced stages. For *EPI* and *HB-EGF*, the distribution between histological types was not equal. *EPI* was elevated in 7% of patients with squamous cell histology (2 of 27) as opposed to 23% (17 of 73) in patients with nonsquamous histology. For *HB-EGF*, the corresponding frequencies were 29% (8 of 27) in squamous cell and 15% (11 of 73) in nonsquamous cell. *HER4* was

Table 2. Characteristics of the EGF System

		Univariate			Multivariate*			Score Value	Nonsquamous ($n = 73$)		Squamous ($n = 27$)	
									Stage I & II ($n = 12$)	Stage III & IV ($n = 61$)	Stage I & II ($n = 7$)	Stage III & IV ($n = 20$)
									N (%)		N (%)	
Biomarker	Cut Point	HR	CI	P Value	HR	CI	P Value	Initial Score	N† (%)	N† (%)	N† (%)	N† (%)
HER1	0.20	0.98	0.48-2.00	.754	1.06	0.65-1.72	.826	−1	1 (8)	24 (39)	3 (42)	5 (25)
HER2	6.40	1.25	0.38-4.09	.707	1.97	0.77-4.99	.155	+1	0	4 (7)	1 (14)	0
HER3	0.12	0.91	0.57-1.46	.710	1.01	0.50-2.07	.959	−1	0	9 (15)	1 (14)	2 (10)
AR	1.03	2.03	0.93-4.44	.077	1.34	0.81-2.22	.254	+1	0	19 (31)	0	6 (30)
EPI	0.25	0.56	0.31-1.02	.060	0.38	0.20-0.72	.003	−1	1 (8)	16 (26)	0	2 (10)
HB-EGF	1.20	1.92	1.11-3.31	.018	1.86	0.96-3.59	.065	+1	0	8 (13)	2 (29)	9 (45)
TGFα	0.20	0.29	0.07-1.20	.086	0.59	0.27-1.32	.203	−1	1 (8)	8 (13)	1 (14)	2 (10)
EGF	0.35	0.22	0.05-0.91	.037	0.99	0.55-1.78	.991	−1	1 (8)	14 (23)	3 (43)	2 (10)
The initial score‡		2.31	1.26-4.56	.008	2.17	1.11-4.22	.02					
The optimized score‡		3.70	2.26-8.08	.00001	2.62	1.49-4.61	.001		0 (0)	12 (20)	1 (14)	8 (40)

* Multivariate results are adjusted for age, sex, PS, stage, and histology.
† Number of patients with upregulated biomarker.
‡ Results are presented with bias-corrected interval.

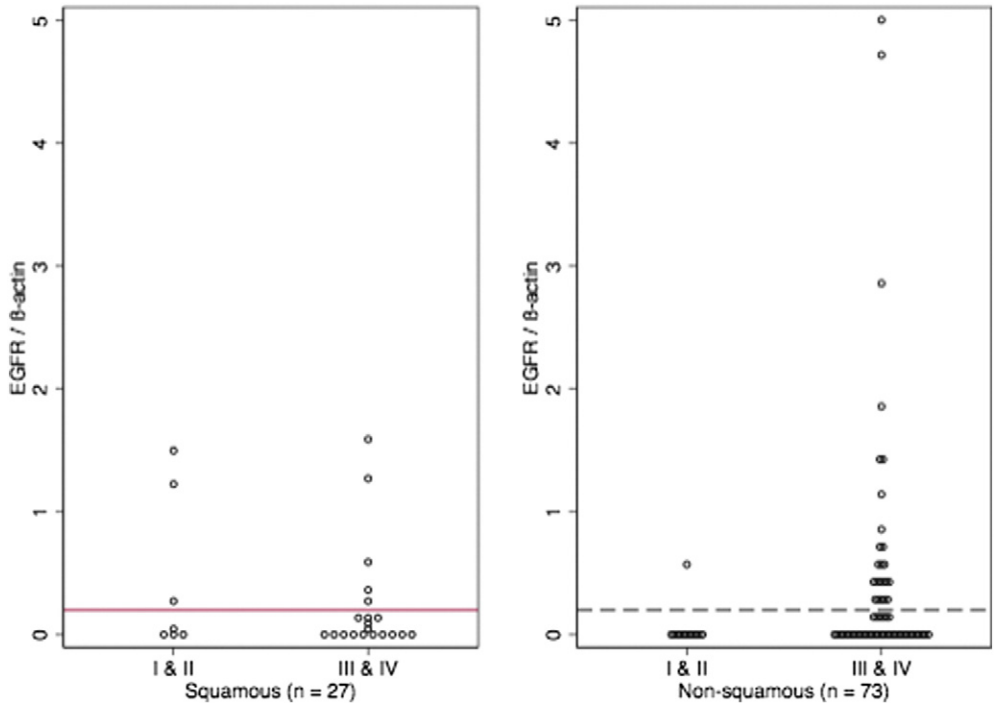

Figure 2. HER1/β-actin levels. Dashed line indicates the cutoff value defined in the reference group. Patients are split by stage: I and II versus III and IV.

hardly measurable in the reference group, and only one patient had elevated *HER4* mRNA. Therefore, *HER4* was not included in further analyses. *HER2* was found in 72 of 100 samples, and yet, only 5 patients had samples with elevated expression compared with the reference group.

2.3. Survival Analyses

In a univariate Cox proportional hazards model, gene expression of all target genes was analyzed for the potential of being an individual prognostic marker (Table 2). This demonstrated that overall *EGFR* gene expression was not a prognostic marker (HR 0.98 [0.48-2.00]; P = .75) as opposed to *EGF* (HR 0.35 [0.05-0.91]; P = .037) and *HB-EGF* (HR 1.2 [1.11-3.31]; P = .018). However, neither of the associations was significant when tested in a multivariate analysis taking known prognostic clinicopathological characteristics into account. For *EPI*, the opposite was found. In a multivariate analysis, high gene expression levels of *EPI* were found to be a marker of better prognosis (HR 0.38 [0.20-0.72]; P = .003). None of the remaining receptors were significant prognostic markers when analyzed individually.

2.4. Score Development

Values creating the initial score are listed in Table 2. The initial score was tested and validated by bootstrapping demonstrating that the score could predict an inferior OS (HR 2.31, BCI [1.26-4.56]). To optimize the initial score, it was tested if one or several of the markers were excessive. Therefore, different combinations were tested in the test cohort. Based on the AIC, the score with the optimal model fit only included the ligands *AREG*, *HBEGF*, *EPI*, and *TGF-α*. In the test cohort as well as in the validation cohort, the optimized score was a prognostic marker of inferior OS (test: HR 6.17, 95% CI [2.60-14.64], P > .001, BCI [1.94-18.89]; validation: HR 2.41, 95% CI [1.23-4.71], P = .01, BCI [1.92-18.61]) (illustrated in Figure 3). In the entire NSCLC group, the optimized score was tested in a multivariate analysis adjusting for the clinical parameters age, sex, PS, stage, and histology. The optimized score was a significant prognostic marker of inferior outcome (HR 2.93 [1.64-5.23], P < .0001, BCI [1.68-4.97]) as seen in Table 3. Median survival for patients with a high score was 4.7 months (95% CI: 2.9-8.3) versus 15.9 months (95% CI: 11.5-22.0) for patients with a low score.

3. Discussion

The impact of *EGFR* expression on OS in NSCLC is controversial. Studies evaluating *EGFR* expression as a prognostic marker have presented contradicting conclusions [13–16]. In a meta-analysis by Nakamura et al., expression of the *EGFR* was not a prognostic marker in NSCLC [13] as opposed to what has been found in other cancers [10–12]. *In vitro* studies have indicated that it is not EGFR in itself but the summed effect of the EGF system that leads to tumor growth [17]. In certain studies, coexpression of the other receptors from the EGF system has been evaluated as a prognostic factor, though without convincing results when known clinical confounders are taken into consideration [21–23]. Expression of the complete network of ligands and receptors interacting with EGFR has not been evaluated as a prognostic factor in NSCLC.

In concordance with the complexity of the EGF system, we have evaluated the gene expression of each receptor and ligand in 100 NSCLC patients *EGFR* wild type in a model. We found that this model predicts the outcome of the individual patient, whereas the individual receptor or ligand does not. Our findings support the hypothesis that it is not the individual receptor or ligand but the network that is reflected in the OS. Gene score models have been used in several settings to sum the effect of several markers that may be coexpressed but differ in impact on a given outcome [19,24,25]. We show by internal validation that our model is a strong prognostic marker in NSCLC patients even when the known clinical characteristics are taken into account. We also demonstrate that it is the increased expression of the ligands *AREG*, *HB-EGF*, *EPI*, and

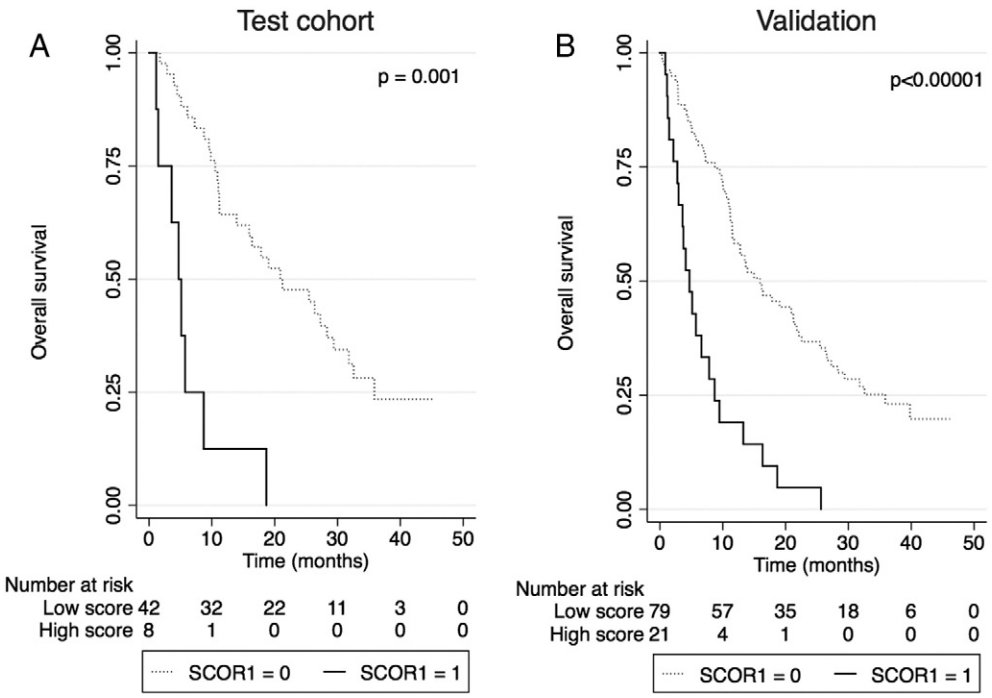

Figure 3. Kaplan-Meier Survival curves for OS in the test cohort (A) and the entire cohort (B). Differences between groups were calculated using the log-rank test.

TGF-α and not the receptors that has the primary impact on the summed effect of the system because removing the effect of the receptors increases the fit of the model. To our knowledge, this has not previously been demonstrated in clinical samples. When establishing the score, we chose an arbitrary approach giving each marker the value of + 1 or − 1 and not the actual coefficient as is done elsewhere [19]. Because of to the small sample size, the actual coefficient for the individual marker is not reliable, especially when splitting into test cohorts of only 50. Optimizing the estimate would lead to a risk of overfitting, and therefore, we chose − 1 or + 1 as coefficients. By doing so, we accept the risk of not being able to incorporate the magnitude of each factor of the EGF system, but in contrast, we do not run the risk of assigning inaccurate coefficients.

A major challenge in explorative gene expression studies is to establish normal values of expression. We identified 111 patients affected by symptoms leading to lung cancer suspicion but where the diagnosis was excluded based on a biopsy from a relevant area of the lung or lymph node. Because of to the consecutive nature of the study, the reference group is younger and there are more never smokers in this group. Despite these differences, we believe the reference group to be valid because it reflects the everyday clinical setting. From the reference group, cut points were defined, resulting in a test where expression above the cut points could be interpreted as truly elevated. Subsequently, we evaluated the prognostic value of the receptors *EGFR*, *HER2,* and *HER3* and the ligands *AREG*, *HB-EGF*, *EPI*, *TGF-α*, and *EGF*. We found that only *EPI* was a prognostic marker of longer OS, whereas none of the receptors or the remaining ligands were significant predictors of outcome in a multivariate analysis taking the clinicopathological characteristics into account. This contradicts other studies [26,27]. In a study by Sunaga et al., it was shown that high *EPI* expression could predict inferior outcome especially in *KRAS*-mutated patients [27]. It is known that 25% to 30% of NSCLC patients are expected to have a mutation in *KRAS*. The *KRAS* mutation status of our patients is unknown, which might explain the contradictory results. Our ability to find a correlation could also be explained by the definition of the cut points. We have used biopsies from patients entering the clinic with symptoms resembling lung cancer to define the cut points, whereas other studies exploring the role of *EPI* have used either the mean or median value of gene expression [27,28] or predefined percentages of cancer cells expressing *EPI*[24,25,29]. By using our approach, we address the fact that the EGF system is not exclusively cancer related but also involved in other biological processes likely to take place in normal tissue (i.e., inflammation, etc.). Furthermore, the samples used in this study were collected with the intent to perform gene expression studies compared with studies where mRNA was extracted from archived formalin-fixed, paraffin-embedded samples.

Despite being unique in this regard, our study faced some limitations. Firstly, the cohort used in this study was consecutively collected, but stages I and II were underrepresented (19 patients), just as only 27 of the patients had squamous histology which was less than expected. We found no effect modification between the score and histology (data not shown), indicating that the expression of ligands had an equal effect in the squamous cell type as well as the nonsquamous *EGFR* wild-type patient. Secondly, though affecting *EGFR*, the receptor *HER4* and the ligand *BTC* were not included in the final score model. *HER4* was

Table 3. Multivariate Cox Proportional Hazard

	HR	95% CI	*P* Value	HR*	95% CI*	*P* Value*	BCI
EGF score	3.52	2.00-6.20	<.001	3.75	2.17-6.47	<.0001	2.09-6.45
Age: ≤67 vs >67	1.28	0.80-2.04	.304				
Sex: female vs male	0.85	0.54-1.32	.462				
PS: 0, 1 vs 2, 3	1.89	1.06-3.37	.030	1.97	1.14-6.47	.014	1.02-4.01
Stage: I & II vs III & IV	4.16	1.92-9.03	<.001	3.63	1.71-7.68	.001	1.47-8.19
Histology: squamous vs nonsquamous	0.72	0.42-1.24	.24				

Results from the multivariate analysis including the optimized EGF score.
* Significant results using backward selection. Removal criterion: $P < .10$.

measurable only in 1 of 50 patients in the test cohort; it was not deemed relevant. In the literature, only a few studies have evaluated *BTC* in clinical samples, and because of limited tissue amount, we chose not to include it in this study. Our findings seem promising, but they need to be validated in an independent cohort. A better understanding of the role of *EGFR* in NSCLC patients without activating mutations in the *EGFR* can help generate new strategies for treatment development and perhaps improve the OS for NSCLC patients.

4. Conclusion

It is well established that activation of *EGFR* is dependent on ligands and other receptors of the EGF system. Therefore, to evaluate *EGFR* as a prognostic marker, it is necessary to take these surrounding ligands and receptors into account. In this study, we show by internal validation that a score combining factors from the EGF system can be developed and that the primary impact on OS is increased expression of the ligands and not the receptors. We have demonstrated a strong prognostic marker that can identify a group of patients where different treatment strategies may be needed.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.tranon.2016.05.002.

Acknowledgement

The authors are thankful to Birgith West Mortensen and Lene Dabelstein Pedersen for laboratory help and to Ninna Aggerholm-Pedersen for statistical help.

References

[1] Siegel R, Ma J, Zou Z, and Jemal A (2014). Cancer statistics. *CA Cancer J Clin* **64**(1), 9–29. http://dx.doi.org/10.3322/caac.21208.

[2] Engholm G, Ferlay J, Christensen N, Kejs AMT, Johannesen TB, Khan S, Leinonen M, Milter MC, Ólafsdóttir E, Petersen T, and Stenz F SH (2011). NORDCAN: Cancer Incidence, Mortality, Prevalence and Survival in the Nordic Countries, Version 7.1 (09.07.2015). Association of the Nordic Cancer RegistriesDanish Cancer Society; 2011 [Available at: http://www.ancr.nu. Accessed December 4, 2015].

[3] American Cancer Society (2014). Cancer Facts and Figures 2014. Available at: http://www.cancer.org/acs/groups/content/@research/documents/webcontent/acspc-042151.pdf; 2014 . [Accessed December 4, 2015].

[4] Cappuzzo F, Hirsch FR, Rossi E, Bartolini S, Ceresoli GL, Bemis L, Haney J, Witta S, Dananberg K, and Domenichini I, et al (2005). Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non–small-cell lung cancer. *J Natl Cancer Inst* **97**(9), 643–655. http://dx.doi.org/10.1093/jnci/dji112.

[5] D'Angelo SP, Janjigian YY, Ahye N, Riely GJ, Chaft JE, Sima CS, Shen R, Zheng J, Dycoco J, and Kris MG, et al (2012). Distinct clinical course of EGFR-mutant resected lung cancers: results of testing of 1118 surgical specimens and effects of adjuvant gefitinib and erlotinib. *J Thorac Oncol* **7**(12), 1815–1822. http://dx.doi.org/10.1097/JTO.0b013e31826bb7b2.

[6] Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, and Haluska FG, et al (2004). Activating mutations in the epidermal growth factor receptor underlying responsiveness of non–small-cell lung cancer to gefitinib. *N Engl J Med* **350**(21), 2129–2139. http://dx.doi.org/10.1056/NEJMoa040938.

[7] Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, and Boggon TJ, et al (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304**(5676), 1497–1500. http://dx.doi.org/10.1126/science.1099314.

[8] Mok TS, Wu Y-L, Thongprasert S, Yang C-H, Chu D-T, Saijo N, Sunpaweravong P, Han B, Margono B, and Ichinose Y, et al (2009). Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* **361**(10), 947–957. http://dx.doi.org/10.1056/NEJMoa0810699.

[9] Zhang Z, Wang T, Zhang J, Zhang Z, Wang T, Zhang J, Cai X, Pan C, Long Y, and Chen J, et al (2014). Prognostic value of epidermal growth factor receptor mutations in resected non–small cell lung cancer: a systematic review with meta-analysis. *PLoS One* **9**(8)e106053. http://dx.doi.org/10.1371/journal.pone.0106053.

[10] Ang KK, Berkey BA, Tu X, Zhang H-Z, Katz R, Hammond EH, Fu KK, and Milas L (2002). Impact of epidermal growth factor receptor expression on survival and pattern of relapse in patients with advanced head and neck carcinoma. *Cancer Res* **62**(24), 7350–7356 [Available at: http://www.ncbi.nlm.nih.gov/pubmed/12499279. Accessed September 29, 2015].

[11] Rokita M, Stec R, Bodnar L, Charkiewicz R, Korniluk J, Smoter M, Cichowicz M, Chyczewski L, Nikliński J, and Kozłowski W (2013). Overexpression of epidermal growth factor receptor as a prognostic factor in colorectal cancer on the basis of the Allred scoring system. *Onco Targets Ther* **6**, 967–976. http://dx.doi.org/10.2147/OTT.S42446.

[12] Inada S, Koto T, Futami K, Arima S, and Iwashita A (1999). Evaluation of malignancy and the prognosis of esophageal cancer based on an immunohistochemical study (p53, E-cadherin, epidermal growth factor receptor). *Surg Today* **29**(6), 493–503 [Available at: http://www.ncbi.nlm.nih.gov/pubmed/10385363. Accessed September 29, 2015].

[13] Nakamura H, Kawasaki N, Taguchi M, and Kabasawa K (2006). Survival impact of epidermal growth factor receptor overexpression in patients with non–small cell lung cancer: a meta-analysis. *Thorax* **61**(2), 140–145. http://dx.doi.org/10.1136/thx.2005.042275.

[14] Timotheadou E, Skarlos D-V, Samantas E, Papadopoulos S, Murray S, Skrickova J, Christodoulou C, Papakostantinou C, Pectasides D, and Papakostas P (2007). Evaluation of the prognostic role of a panel of biomarkers in stage IB-IIIA non–small cell lung cancer patients. *Anticancer Res* **27**(6C), 4481–4489 [Available at: http://www.ncbi.nlm.nih.gov/pubmed/18214064. Accessed September 1, 2015].

[15] Seyhan EC, Altın S, Cetinkaya E, Sökücü S, Abalı H, Buyukpinarbasili N, Fener N, Kaplanova J, and Vrettou E (2010). Prognostic value of epidermal growth factor receptor expression in operable non–small cell lung carcinoma. *Multidiscip Respir Med* **5**(5), 305–311. http://dx.doi.org/10.1186/2049-6958-5-5-305.

[16] Van Dyke AL, Cote ML, Prysak GM, Claeys GB, Wenzlaff AS, Murphy VC, Lonardo F, and Schwartz AG (2008). COX-2/EGFR expression and survival among women with adenocarcinoma of the lung. *Carcinogenesis* **29**(9), 1781–1787. http://dx.doi.org/10.1093/carcin/bgn107.

[17] Citri A and Yarden Y (2006). EGF-ERBB signalling: towards the systems level. *Nat Rev Mol Cell Biol* **7**(7), 505–516. http://dx.doi.org/10.1038/nrm1962.

[18] Sandfeld-Paulsen B, Demuth C, Folkersen BH, Rasmussen TR, Madsen LB, Sorensen BS, and Meldgaard P (2016). Ultra-micro samples can be used for mRNA quantification of lung cancer biomarkers. *Scand J Clin Lab Invest* . http://dx.doi.org/10.3109/00365513.2016.1143114.

[19] Chen H-Y, Yu S-L, Chen C-H, Chang G-C, Chen C-Y, Yuan A, Cheng C-L, Wang C-H, Terng H-J, and Kan S-F, et al (2007). A five-gene signature and clinical outcome in non–small-cell lung cancer. *N Engl J Med* **356**(1), 11–20. http://dx.doi.org/10.1056/NEJMoa060096.

[20] Burnham K (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. . 2nd ed.New York: Springer; 2002 .

[21] Brabender J, Danenberg KD, Metzger R, Schneider PM, Park J, Salonga D, Hölscher AH, and Danenberg PV (2001). Epidermal growth factor receptor and HER2-neu mRNA expression in non–small cell lung cancer is correlated with survival. *Clin Cancer Res* **7**(7), 1850–1855 [Available at: http://www.ncbi.nlm.nih.gov/pubmed/11448895. Accessed September 1, 2015].

[22] Onn A, Correa AM, Gilcrease M, Isobe T, Massarelli E, Bucana CD, O'Reilly MS, Hong WK, Fidler IJ, and Putnam JB, et al (2004). Synchronous overexpression of epidermal growth factor receptor and HER2-neu protein is a predictor of poor outcome in patients with stage I non–small cell lung cancer. *Clin Cancer Res* **10**(1 Pt 1), 136–143 [Available at: http://www.ncbi.nlm.nih.gov/pubmed/14734462. Accessed September 2, 2015].

[23] Lai WW, Chen FF, Wu MH, Chow NH, Su WC, Ma MC, Su PF, Chen H, Lin MY, and Tseng YL (2001). Immunohistochemical analysis of epidermal growth factor receptor family members in stage I non–small cell lung cancer. *Ann Thorac Surg* **72**(6), 1868–1876 [Available at: http://www.ncbi.nlm.nih.gov/pubmed/11789762. Accessed November 27, 2015].

[24] Subramanian J and Simon R (2010). Gene expression-based prognostic signatures in lung cancer: ready for clinical use? *J Natl Cancer Inst* **102**(7), 464–474. http://dx.doi.org/10.1093/jnci/djq025.

[25] Harris BHL, Barberis A, West CML, and Buffa FM (2015). Gene expression signatures as biomarkers of tumour hypoxia. *Clin Oncol* **27**(10), 547–560. http://dx.doi.org/10.1016/j.clon.2015.07.004.

[26] Zhang J, Iwanaga K, Choi KC, Wislez M, Raso MG, Wei W, Wistuba II, and Kurie JM (2008). Intratumoral epiregulin is a marker of advanced disease in non–small cell lung cancer patients and confers invasive properties on EGFR-mutant cells. *Cancer Prev Res (Phila)* **1**(3), 201–207. http://dx.doi.org/10.1158/1940-6207.CAPR-08-0014.

[27] Sunaga N, Kaira K, Imai H, Shimizu K, Nakano T, Shames DS, Girard L, Soh J, Sato M, and Iwasaki Y, et al (2013). Oncogenic KRAS-induced epiregulin overexpression contributes to aggressive phenotype and is a promising therapeutic target in non–small-cell lung cancer. *Oncogene* **32**(34), 4034–4042. http://dx.doi.org/10.1038/onc.2012.402.

[28] Koeppen H, Yu W, Zha J, Pandita A, Penuel E, Rangell L, Raja R, Mohan S, Patel R, and Desai R (2014). Biomarker analyses from a placebo-controlled phase II study evaluating erlotinib±onartuzumab in advanced non–small cell lung cancer: MET expression levels are predictive of patient benefit. *Clin Cancer Res* **20**(17), 4488–4498. http://dx.doi.org/10.1158/1078-0432.CCR-13-1836.

[29] Sun M, Behrens C, Feng L, Ozburn N, Tang X, Yin G, Komaki R, Varella-Garcia M, Hong WK, and Aldape KD, et al (2009). HER family receptor abnormalities in lung cancer brain metastases and corresponding primary tumors. *Clin Cancer Res* **15**(15), 4829–4837. http://dx.doi.org/10.1158/1078-0432.CCR-08-2921.